UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/828,281	04/21/2004	Ji Hyun Wang	123056-05004422	6375

22429   7590   08/22/2007

LOWE HAUPTMAN HAM & BERNER, LLP
1700 DIAGONAL ROAD
SUITE 300
ALEXANDRIA, VA 22314

EXAMINER
HERNANDEZ, JOSIAH J

ART UNIT	PAPER NUMBER
2626	

MAIL DATE	DELIVERY MODE
08/22/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

PTOL-90A (Rev. 04/07)

	Application No.	Applicant(s)
Office Action Summary	10/828,281	WANG ET AL.
	Examiner	Art Unit
	Josiah Hernandez	2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *21 April 2004*.

2a)☐ This action is **FINAL**.   2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-14* is/are pending in the application.

    4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-14* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

    a)☒ All b)☐ Some * c)☐ None of:

      1.☒ Certified copies of the priority documents have been received.

      2.☐ Certified copies of the priority documents have been received in Application No. _____.

      3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

    * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☒ Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date *04/21/2004*.

4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____ .

5) ☐ Notice of Informal Patent Application

6) ☐ Other: _____.

DETAILED ACTION

Claim Rejections - 35 USC § 103

1.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set
> forth in section 102 of this title, if the differences between the subject matter sought to be patented and
> the prior art are such that the subject matter as a whole would have been obvious at the time the
> invention was made to a person having ordinary skill in the art to which said subject matter pertains.
> Patentability shall not be negatived by the manner in which the invention was made.

2.   Claims 1 and 14 are rejected under 35 U.S.C. 103(a) as being unpatentable over

Kupiec (US 5,519,608) in view of Saito et al. (US PGPub 2001/0042083).

As to claim 1, Kupiec discloses a semi-automatic contruction method for

knowledge base (abstract lines 1-5) of an encyclopedia question answering

system (column 9 lines 10-15), the method comprising the steps of: extracting

unstructured information from a body of the encyclopedia (column 9 lines 5-10).

Kupiec does not disclose specifically designing structure of knowledge

base using templates, extracting structured information and storing the

information in the templates. Saito teaches using templates for extracting

information from documents (title) of which designs the structure of the

knowledge base with a plurality of templates for each entry and a plurality of

attributes related to each of the templates (for each entry or document a user can

specify a template and its attributes by giving labels for selected areas of

extraction and characteristic extraction units, paragraphs [0006] lines 7-12 and [0007] lines 7-16). Saito also teaches extracting structured (paragraph [0055] lines 10-15) and unstructured (paragraph [0056] lines 1-10) information including an attribute name and value of the entry from a summary or body of the text (paragraphs [0006] lines 7-12 and [0007] lines 7-16); and storing the structured information and the unstructured information in corresponding template and attribute of the knowledge base according to the entry (the template is specified and its structure and attributes are stored and used to extract information and store that in the template form in a database, paragraph [0037] lines 10-22).

It would have been obvious to one having ordinary skill in the art at the time the invention was made to have modified the method of Kupiec with the template, extraction, and storing of information and its attributes. Doing so would have allowed to provide a more systematic and flexible method to extract information out of encyclopedia documents. The templates would have allowed a higher degree of structure to the searching and extraction method.

As to claim 14, Kupiec does not disclose specifically storing attribute information. Saito teaches constructing the knowledge base with the attribute name and the attribute values extracted (this is done by setting the attribute information of the templates and extracting information according to the templates, paragraphs [0006] lines 7-12 and [0007] lines 7-16) and additionally storing the attribute name and the attribute values extracted as the unstructured

information in the knowledge base according to existence of same attribute value of the entry (paragraph [0051] lines 17-27).

It would have been obvious to one having ordinary skill in the art at the time the invention was made to have modified the method of Kupiec with the structure of Saito. Doing so would have allowed for efficient method to organize multiple attribute values so that they are not confused with the other similar attributes.

3.   Claims 2-6 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kupiec (US 5,519,608) in view of Saito et al. (US PGPub 2001/0042083) as applied in claim 1 and in further view of Tan et al. (US PGPub 2006/0026203).

As to claim 2, Saito teaches individual attribute templates of a specific attribute of an individual category of the encyclopedia, for each entry (paragraphs [0006] lines 7-12 and [0007] lines 7-16). Kupiec or Saito do not specifically disclose constructing knowledge base with common attribute templates. Tan teaches a method for discovering knowledge from text (title) and the structure of the knowledge base is constructed with common attribute templates of a common attribute shared in categories of the text (abstract lines 1-7).

It would have been obvious to one having ordinary skill in the art at the time the invention was made to have modified the method of Kupiec with the

structure of Saito and Tan. Doing so would have allowed for the formation of the

knowledge base in a much more organized fashion with out having to create

numerous templates and reducing the amount by grouping the similar attributes

into the templates.

As to claim 3, Kupiec or Saito do not specifically disclose using attributes

having similar meanings. Tan teaches attributes having similar meaning are

managed as a representative attribute integrally and detail meanings of the

attributes are grouped and defined in separate subgroup fields (paragraph

[0010]).

It would have been obvious to one having ordinary skill in the art at the

time the invention was made to have modified the method of Kupiec with the

structure of Tan. Doing so would have allowed for the formation of the

knowledge base in a much more organized fashion with out having to create

numerous templates and reducing the amount by grouping the similar attributes

into the templates.

As to claim 4, Saito discloses extracting the engry, the attribute name and

the attribute values (for each entry or document a user can specify a template

and its attributes by giving labels for selected areas of extraction and

characteristic extraction units, paragraphs [0006] lines 7-12 and [0007] lines 7-

16).

Kupiec or Saito do not specifically disclose recognizing patterned format. Tan teaches recognizing a patterned format of the summary information (abstract lines 11-17).

It would have been obvious to one having ordinary skill in the art at the time the invention was made to have modified the method of Kupiec and Saito with the structure of Tan. Doing so would have allowed for efficient searching for structured summary information.

As to claim 5, Saito discloses only if the attribute name belongs to the valid attribute in the attribute list of the templates of the knowledge base, extracting the corresponding attribute value (for each entry or document a user can specify a template and its attributes by giving labels for selected areas of extraction and characteristic extraction units, paragraphs [0006] lines 7-12 and [0007] lines 7-16 and paragraphs [0006] lines 7-12 and [0007] lines 7-16) Kuipec or Saito do not specifically disclose using patterned format and determining if the extracted information belongs to the pattern. Tan teaches extracting the entry and the attribute name through the patterned format of the summary information (abstract lines 11-17); ascertaining whether the attribute name belongs to a valid attribute in an attribute list of the templates of the knowledge base (once the pattern formats are used the extracted text are chosen according to the patterns, paragraph 12-14).

It would have been obvious to one having ordinary skill in the art at the time the invention was made to have modified the method of Kupiec and Saito with the structure of Tan. Doing so would have allowed for efficient searching for structured summary information.

As to claim 6, Kupiec does not specifically disclose using marked identifiers. Saito teaches if the extracted attribute name has a plurality of attribute values, extracting each of the plurality of attribute values separately by marked identifier (paragraph [0051] lines 17-27).

It would have been obvious to one having ordinary skill in the art at the time the invention was made to have modified the method of Kupiec and Saito with the structure of Tan. Doing so would have allowed for efficient method to organize multiple attribute values so that they are not confused with the other similar attributes.

4.   Claims 7-13 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kupiec (US 5,519,608) in view of Saito et al. (US PGPub 2001/0042083) as applied in claim 1 and in further view of Paik et al. (US 6,263,335) and McCarley (US PGPub 2004/0122656).

As to claim 7, Kupiec or Saito do not disclose specifically using token strings and dependence relations. Paik teaches converting each sentence of an illustrative corpus into a token string, recognizing dependence relation of an attribute tagging token, generating learning data, and learning the learning data through a predetermined model; and converting each sentence of the body of the encyclopedia into the token string, recognizing dependence relation of an extraction object tokens, and applying a learning result and the model to a recognition result, thereby finding and extracting the attribute name and the attribute value of each extraction object token (a concept-relation-concept triples is used to convert the text into knowledge representation (tokens) then dependence relations is recognized generating data for a model, column 3 lines 59-67 and column 4 lines 1-5). Paik does not disclose specifically using stochastic models. McCarley teaches a system for understanding documents and text (abstract lines 1-7) and the use of maximum entropy model as the stochastic model (column 4 lines 50-60).

It would have been obvious to one having ordinary skill in the art at the time the invention was made to have modified the method of Kupiec and Saito with the use of stochastic models as taught by McCarley and the use of token strings and dependence relations as taught by Paik. Doing so would have allowed high accuracy in NLU systems when using statistical models.

As to claim 8, Kupiec or Saito do not disclose specifically morpheme

parsing. Paik teaches performing morpheme parsing on the illustrative corpus of

the encyclopedia (column 12 lines 53-54 and column 9 lines 53-60), which is

tagged with an object name and an attribute, and recognizing a word phrase unit

token string according to sentences (column 3 lines 59-67 and column 4 lines 1-

5); applying a predetermined dependence rule to a token tagged with an attribute

value in the token string, thereby recognizing dependence for the object token;

and generating the learning data by using the governor and the dependent of

each object token as contexts, and storing the learning result in the stochastic

model (a concept-relation-concept triples is used to convert the text into

knowledge representation (tokens) then dependence relations is recognized

generating data for a model, column 3 lines 59-67 and column 4 lines 1-5). Paik

does not disclose specifically using stochastic models. McCarley teaches a

system for understanding documents and text (abstract lines 1-7) and the use of

maximum entropy model as the stochastic model (column 4 lines 50-60).

It would have been obvious to one having ordinary skill in the art at the

time the invention was made to have modified the method of Kupiec and Saito

with the use of stochastic models as taught by McCarley and the use of parsing

as taught by Paik. Doing so would have allowed high accuracy in NLU systems

when using statistical models.

As to claim 9, Kupiec or Saito do not disclose specifically morpheme parsing. Paik teaches performing the morpheme parsing and object name recognition (column 12 lines 53-54 and column 9 lines 53-60) on the body of the encyclopedia, and converting each sentence into the word-phrase unit token string; designating a token of the token string as an extraction object token, the token of the token string having object name or full morpheme as a noun (column 6 lines 9-13, column 3 lines 59-67, and column 4 lines 1-5); applying a predetermined dependence rule to each of the designated extraction object tokens, and recognizing a context token of the governor and the dependent; and applying the extraction object token and the context token to the learning result and the stochastic model, grouping attribute types of the extraction object tokens, and extracting the attribute type of the extraction object tokens that have highest probabilities with the attribute names and the attribute values of the extraction object token (a concept-relation-concept triples is used to convert the text into knowledge representation (tokens) then dependence relations is recognized generating data for a model, column 3 lines 59-67 and column 4 lines 1-5). Paik does not disclose specifically using stochastic models. McCarley teaches a system for understanding documents and text (abstract lines 1-7) and the use of maximum entropy model as the stochastic model (column 4 lines 50-60).

It would have been obvious to one having ordinary skill in the art at the time the invention was made to have modified the method of Kupiec and Saito

with the use of stochastic models as taught by McCarley and the use of parsing as taught by Paik. Doing so would have allowed high accuracy in NLU systems when using statistical models.

As to claim 10, Kupiec or Saito do not disclose specifically the use of verb phrases, adverbial cases, and nouns. Paik teaches the dependence rule used to recognize the dependence relation, the governor is a verb phrase nearest to the dependent (column 6 lines 48-51) if the dependence is any one selected from the group consisting of a subjective case, and objective case and an adverbial case (column 12 lines 40-48), and the governor is a noun nearest to the dependent if the dependent is any one selected from the group consisting of an adnorminal phrase and an adnorminal clause (column 6 lines 48-51).

It would have been obvious to one having ordinary skill in the art at the time the invention was made to have modified the method of Kupiec and Saito with the use of dependence rule as taught by Paik. Doing so would have allowed to accurately identify the relationship between words in order to understand context and meaning from a text.

As to claim 11, Kupiec or Saito do not disclose specifically the use of neighboring nouns. Paik teaches the dependence rule used to recognize the dependence relation, in case of neighboring nouns and/or object names (column 3 lines 65-67 and column 4 lines 1-5), a preceding noun or a preceding object

name is a dependent and a following noun or a following object name is a governor (column 6 lines 48-51).

It would have been obvious to one having ordinary skill in the art at the time the invention was made to have modified the method of Kupiec and Saito with the use of dependence rule as taught by Paik. Doing so would have allowed to accurately identify the relationship between words in order to understand context and meaning from a text.

As to claim 12, Kupiec or Saito do not disclose specifically the use of tokens around an object name. Paik teaches the dependence rule used to recognize the dependence relation, when tokens around an object name or nouns are symbols, a verb phrase of a sentence is a governor (column 6 lines 48-51).

It would have been obvious to one having ordinary skill in the art at the time the invention was made to have modified the method of Kupiec and Saito with the use of dependence rule as taught by Paik. Doing so would have allowed to accurately identify the relationship between words in order to understand context and meaning from a text.

As to claim 13, Kupiec or Saito do not disclose specifically using a maximum entropy model as the stochastic model. McCarley teaches a system for understanding documents and text (abstract lines 1-7) and the use of maximum entropy model as the stochastic model (column 4 lines 50-60).

It would have been obvious to one having ordinary skill in the art at the time the invention was made to have modified the method of Kupiec and Saito with the use of maximum entropy models as taught by McCarley. Doing so would have allowed high accuracy in NLU systems when using a stochastic model.

Conclusion

Any inquiry concerning this communication should be directed to Josiah Hernandez whose telephone number is 571-270-1646. The examiner can normally be reached from 7:30 pm to 5:00 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David Hudspeth can be reached on (571) 272-7843. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from

the Patent Application Information Retrieval (PAIR) system. Status information

for published applications may be obtained from either Private PAIR or Public

PAIR. Status information for unpublished applications is available through

Private PAIR only. For more information about the PAIR system, see http://pair-

direct.uspto.gov. Should you have questions on access to the Private PAIR

system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-

free).

DAVID HUDSPETH
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER

JH